Traditional HPC needs: particle accelerators

Jean-Luc Vay

With inputs from

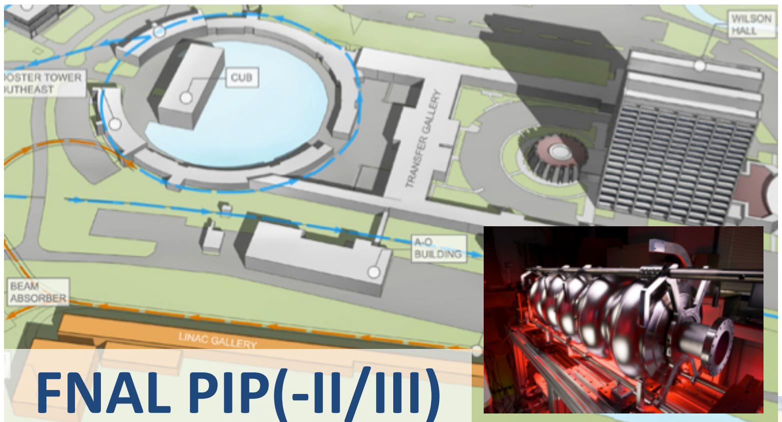
J. Amundson, J. Cary, W. Mori, C.-K. Ng, R. Ryne, J. Qiang

Exascale Requirements Reviews:
High Energy Physics

June 10-12, 2015

Advanced simulations play an increasingly important role in the design, operation and tuning of accelerators.

"Conventional accelerators"
accelerate beams in RF cavities

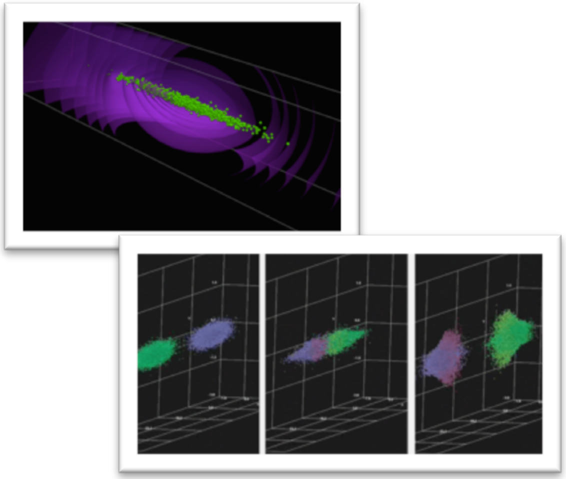


"Advanced concepts"
accelerate beams in plasmas

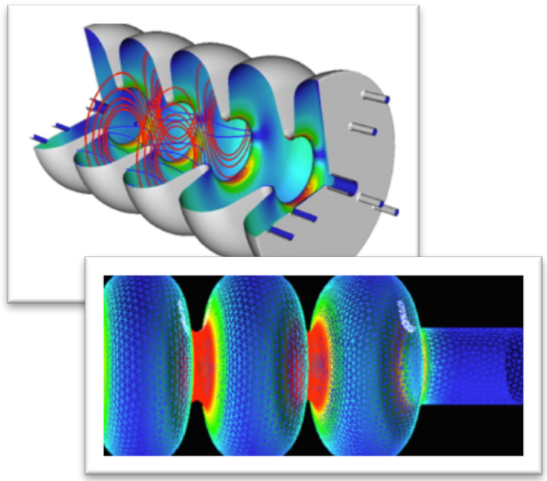
Ultimate goal: real-time virtual prototyping of entire accelerators.

Large scale modeling split along three categories

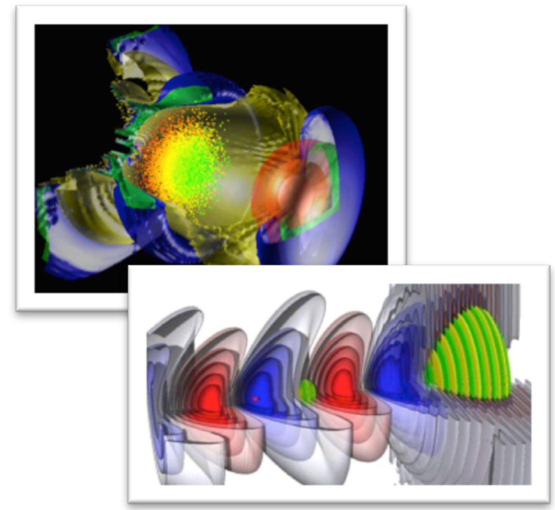
Beam dynamics
Generation & transport in external+self-fields



Electromagnetics
Prediction of fields in components (RF)



Plasma wakefield
Self-consistent evolution of wake and beam



Computational methods based on:

- **Particle-In-Cell** approach (electrostatic, electromagnetic, quasi-static),
- on **structured finite-difference** or **unstructured finite-elements** grids,
- parallelism by **domain decomposition** using **MPI** (all), MPI+OpenMP/Cuda (some).

DOE-HEP funds fully or partially the development of a number of parallel accelerator simulation tools

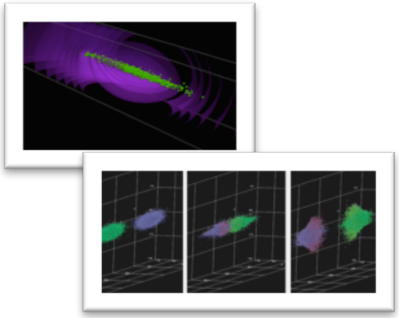
Name	Beam dynamics	Electro-magnetics	Plasma accelerator	Institution	Distribution	Methods FE=Finite Element; FD=Finite Difference ES=electrostatic; EM=electromagnetic QS=Quasistatic; PS=Pseudo-Spectral (FFT)		
ACE3P		X		SLAC	Upon request	FE	EM	PIC
BeamBeam3D	X			LBNL	Upon request	FD	ES	PIC
IMPACT	X			LBNL	Upon request	FD	ES	PIC
Osiris			X	UCLA	Upon request	FD	EM	PIC
QuickPIC			X	UCLA	Upon request	FD	QS	PIC
Synergia	X			FNAL	Open source	FD	ES	PIC
UPIC			X	UCLA	Upon request	PS	EM	PIC
Vorpal		X	X	Tech-X	Commercial	FD	EM	PIC
Warp	X		X	LBNL/LLNL	Open source	FD/PS	ES/EM	PIC

All are – or have been – funded in part by SciDAC.

Languages: FORTRAN, C, C++. Some use Python as frontend.

Challenges will require exascale supercomputing

Beam dynamics



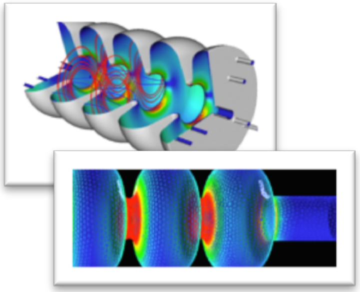
Now: 10^6-10^7 CPU-hours

2025: 10^8-10^{10} CPU-hours

For modeling of beam and/or beam-beam crossing in PIP-II/III/HL-LHC:

Requires $10^5 - 10^8$ steps through:
- 1,000s of elements
- 1,000s-1,000,000s of revolutions
- 1-1000s bunches of O(10^{12}) part.
- 100s of cases

Electromagnetics



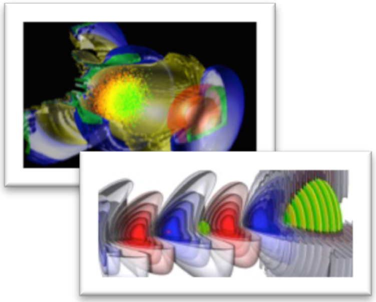
Now: 10^6-10^7 CPU-hours

2025: 10^8-10^{10} CPU-hours

For modeling of dark currents & radiation effects in PIP-II:

Requires $10^5 - 10^8$ steps through :
- O(10^8) macro-particles
- O(10^{10}-10^{12}) grid elements
- 25 cryomodules
- 10s-100s of cases

Plasma wakefield



Now: 10^6-10^7 CPU-hours

2025: 10^8-10^{12} CPU-hours

For modeling of plasma-based collider (beam or laser driven):

Requires $10^4 - 10^7$ steps through :
- O(10^{10}-10^{12}) grid cells/stage
- O(10^{12}) plasma macro-part./stage
- 20-100 stages
- 10s-100s cases

Parameter scans are crucial for optimization of accelerator design.

Exascale will impose significant developments

Accelerator codes are transitioning to **deeper multi-level parallelism**:

- intra-core: vectorization
- **intra-node: multi-threaded openMP/CUDA (newest)**
- inter-nodes: MPI (can be multi-level: e.g. multi-bunch with 1/bunch/MPI group)
- ensemble: parallel optimization

Scaling to large # of cores will necessitate **efficient dynamic load balancing**

- scaling of EM-PIC with uniform plasma or ES-PIC multi-bunch is easy
- general case of non-uniform plasma, intra-bunch is much harder

A **broad range of field solvers need to scale** to large # of cores

- 2^{nd} order FD, high-order FD & FE, spectral (FFT), multi-grid, linear algebra
- adaptive mesh refinement (not mainstream and will need extra development)

Finite-element electromagnetics can be **memory intensive**: currently ~64G/node

High-volume of data increases the need for **efficient parallel IO/in-situ analysis/viz.**

Exascale is an opportunity for the accelerator community

HEPAP subpanel report:

- "Over the past 75 years, ***accelerator science*** and ***technology*** has contributed to research that led to ***twenty-five Nobel Prizes in Physics***."
- "***Computer simulations*** play an ***indispensable role*** in ***all accelerator areas***."
- ***"Advancing the capabilities of accelerator simulation codes*** to capitalize on the ***drive toward exascale computing*** would have ***large benefits*** in improving accelerator design and performance."
- ***"Advanced simulation tools will maximize the productivity of R&D for all future accelerators***."

Coupled to algorithmic advances, it will enable reaching the ultimate goal of real-time virtual prototyping of entire accelerators.

Thank you.

Extras

Case studies

Beam dynamics
BeamBeam3D

Code: ___BeamBeam3D___________	Column 1: Current Usage	Future Usage: 2020 (As a factor of column 1)****	Future Usage: 2025 2 (As a factor of column 1)****
Computational core hours (Conventional)*	1.5 million	15 million	100 million
Computational node hours (Homogeneous many-core)**			
Computational node hours (w/GPU or accelerator)***		3 million	6 million
Memory per node	1 GB	1 GB	1 GB
Aggregate memory	0.1 TB	10 TB	100 TB
Data read and written per run	0.1 TB	1 TB	10 TB
Maximum I/O bandwidth needed	10 GB/sec	10 GB/sec	10 GB/sec
Percent of runtime for I/O	<5%	<5%	5%
Scratch file system space needed	1 TB	10 TB	20 TB
Permanent online data storage	2 TB	20 TB	40 TB
Archival data storage needed	4 TB	50 TB	100 TB

Electromagnetics
ACE3P

Code: ACE3P	Column 1: Current Usage	Future Usage: 2020 (As a factor of column 1)****	Future Usage: 2025 2 (As a factor of column 1)****
Computational core hours (Conventional)*	2.5 M	15 M	50 M
Computational node hours (Homogeneous many-core)**			
Computational node hours (w/GPU or accelerator)***			
Memory per node	64 GB	96 GB	128 GB
Aggregate memory	2 TB	20 TB	40 TB
Data read and written per run	2 TB	10 TB	20 TB
Maximum I/O bandwidth needed	20 GB/sec	40 GB/sec	80 GB/sec
Percent of runtime for I/O	20	20	20
Scratch file system space needed	50 TB	200 TB	500 TB
Permanent online data storage	5 TB	20 TB	50 TB
Archival data storage needed	100 TB	400 TB	800 TB

Case study: BeamBeam3D

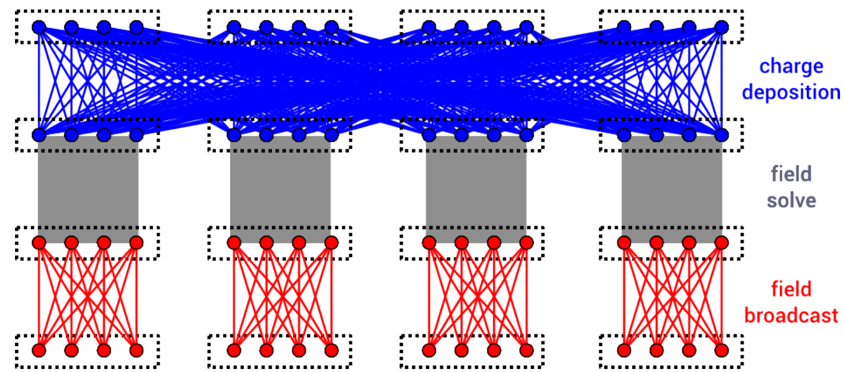
Code: ___BeamBeam3D___________	**Column 1: Current Usage**	**Future Usage: 2020 (As a factor of column 1)******	**Future Usage: 2025 2 (As a factor of column 1)******
Computational core hours (Conventional)*	1.5 million	15 million	100 million
Computational node hours (Homogeneous many-core)**			
Computational node hours (w/GPU or accelerator)***		3 million	6 million
Memory per node	1 GB	1 GB	1 GB
Aggregate memory	0.1 TB	10 TB	100 TB
Data read and written per run	0.1 TB	1 TB	10 TB
Maximum I/O bandwidth needed	10 GB/sec	10 GB/sec	10 GB/sec
Percent of runtime for I/O	<5%	<5%	5%
Scratch file system space needed	1 TB	10 TB	20 TB
Permanent online data storage	2 TB	20 TB	40 TB
Archival data storage needed	4 TB	50 TB	100 TB

Case study: ACE3P

Code: ACE3P	**Column 1: Current Usage**	**Future Usage: 2020 (As a factor of column 1)******	**Future Usage: 2025 2 (As a factor of column 1)******
Computational core hours (Conventional)*	2.5 M	15 M	50 M
Computational node hours (Homogeneous many-core)**			
Computational node hours (w/GPU or accelerator)***			
Memory per node	64 GB	96 GB	128 GB
Aggregate memory	2 TB	20 TB	40 TB
Data read and written per run	2 TB	10 TB	20 TB
Maximum I/O bandwidth needed	20 GB/sec	40 GB/sec	80 GB/sec
Percent of runtime for I/O	20	20	20
Scratch file system space needed	50 TB	200 TB	500 TB
Permanent online data storage	5 TB	20 TB	50 TB
Archival data storage needed	100 TB	400 TB	800 TB

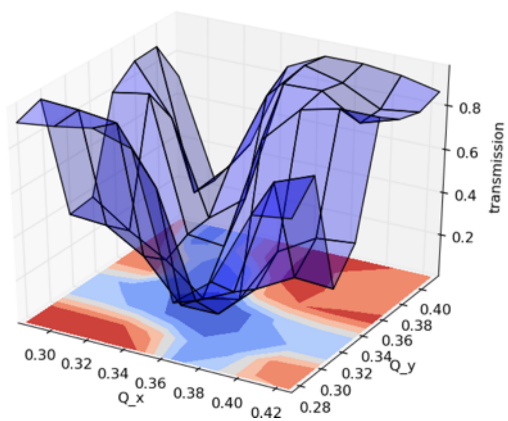
Scaling for future Synergia needs

- Scaling advances are the product of many factors
  - Redundant solves (communication avoidance) (x4-x10)
    - Every simulation
  - Large statistics (x1-x1000)
    - Some simulations
  - Multiple bunches (x1-x1000)
    - Some simulations
  - Parameter scans (x1 – x100)
    - Some simulations
- Product can be huge (x4 – x1e8)



Communication avoidance avoids limiting scaling to the least scalable element

Multiple bunch simulations Scale as O(1) in the number of bunches

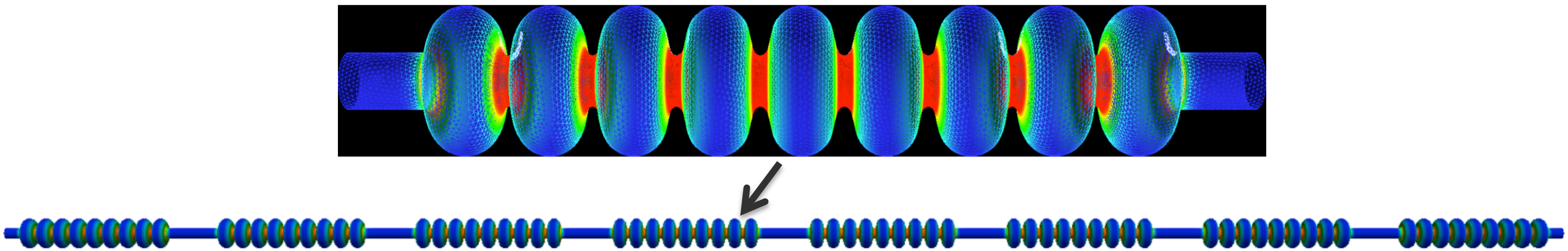


Parameter scans are crucial for accelerator design

Fermilab

ACE3P: End-to-End Simulation of Dark Current and Radiation

- **Cryomodule simulation**



Dark current simulation for a single LCLS-II cromodule

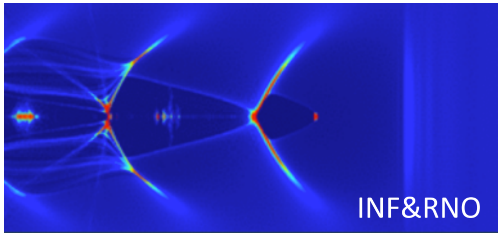
– 30 million particles, 20 minutes using 4800 cores on NERSC Edison

- **Linac simulation**
  - Increase in problem size from modeling multiple cyomodules in superconducting linacs
    - 25 cryomodules of different cavity designs in PIP-II
    - 30 cryomodules in LCLS-II
  - Integrated simulation of dark current and radiation effects (with Geant4) increases number of particles from secondary electrons.
  - 2-3 orders more computing resources required

Large scale simulations take too long! Real machines need fast turnaround for real-time tuning.

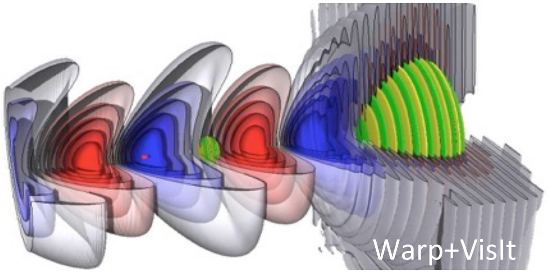
Now

2D-RZ simulations of BELLA



0.5 M CPU-HRS (~1 week for 9 cm plasma)

3D simulation of novel injection scheme



0.25 M CPU-HRS (~1.5 day for 1 mm plasma)

10-year projection

- 3-D BELLA simulation: **~ 500 M CPU-HRs**
- Parametric studies: **× 100 cases**
- Collider: **× 100 stages**

Trillions of CPU-HRs

Algorithm advances + exascale are a necessity!

*W. P. Leemans, *et al.*, *Phys. Rev. Lett.* **113**, 245002 (2014)
L.-L. Yu, *et al.*, *Phys. Rev. Lett.* **112, 125001 (2014)

Requirements for a single PWFA stage: A collider will have ~20 stages

25 GeV Stage

Drive Beam : σ_r = 3.3 μm , σ_z = 30.0 μm , N_1 = 3.0 x 10^{10} , ε = 100 mm mrad
Trailing Beam: σ_r = 103.7 nm, σ_z = 10.0 μm, N_2 = 1.0 x 10^{10}, ε = 0.1 mm mrad
Distance between two beams : 115 μm; Plasma Density : 1.0 x 10^{17} cm^{-3}; Plasma Length: 1 m

QuickPIC Simulation of PWFA Linear Collider

Simulation Parameters

Upper Limit

Size: 400 μm x 400 μm x 300 μm; Cells: 32,768 x 32,768 x 2048
Simulated Beam Particles: 1.0 x 10^{10}; Simulated Plasma Particles (electrons and ions): 6.4 x10^{10}
Total Time Steps: 1.2 x 10^4; Estimated CPU-Hours: 8.8 x 10^8 (2 us per particle per 2D step)

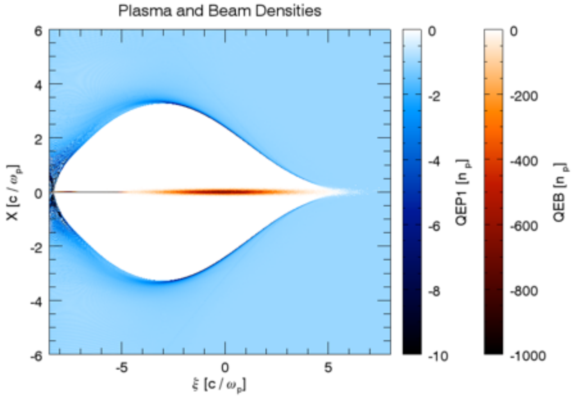
Lower Limit

Size: 400 μm x 400 μm x 300 μm; Cells: 4096 x 4096 x 512
Simulated Beam Particles: 1.0 x 10^8; Simulated Plasma Particles (electrons): 6.7 x 10^7
Total Time Steps: 3 x 10^3; Estimated CPU-Hours: 5.7 x 10^4 (2 ms per particle per 2D step)

20 Stages

Upper Limit: 1.8 x 10^{10} CPU-Hours
Lower Limit: 1.1 x 10^5 CPU-Hours

Based on a single predictor-corrector loop



QuickPIC simulation of a 25 GeV PWFA stage for accelerating electrons.